

Data Fusion in Wireless Sensor Networks

Maen Takruri

Submitted in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

Faculty of Engineering and Information Technology
UNIVERSITY OF TECHNOLOGY, SYDNEY


March 2009

Copyright © 2009 Maen Takruri

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm or any other means without written permission from the author.

Certificate of Authorship/Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text. I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

A handwritten signature in dark ink, appearing to read 'Maen Takruri', is written over a horizontal line.

Maen Takruri, March 2009

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Subhash Challa, whose generosity and commitment are above and beyond the call of duty; words do not describe my gratitude. I appreciate his vast knowledge and skill in many areas and his assistance in completing this thesis. I would like also to thank my co-supervisor, Dr. Tim Aubrey for his assistance and support.

I must also thank the following people from the University of Technology, Dr. Khalid Aboura for his assistance and advice in statistics, Dr. Rami Al-Hmouz, Dr. Mohammad Momani, Dr. Kais Al-Momani, Mr. Mohammad Al-Hattab and Mr. Akram AlSukker for their support and invaluable philosophical debates, exchanges of knowledge, which helped enrich the experience.

I would like to acknowledge the support of the ARC Research Network on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) through the collaborative work with A/Prof. Marimuthu Palaniswami, A/Prof. Christopher Leckie, Mr. Sutharshan Rajasegarar from the University of Melbourne, to whom I would like to express my deep and sincere gratitude for their valuable contribution in implementing Support Vector Regression with the Drift detection and correction algorithms in chapters 6 and 7 of this thesis. Thanks also go to Dr. Rajib Chakravorty from NICTA Victoria Laboratory for his assistance and advice in estimation theory.

I am deeply grateful to my sisters Ruba, Heba and Sahar and my brother Awn, for their loving support. I owe my loving thanks to my wife Ramah and my beloved daughter Tasneem. Without their love, patience and encouragements I would not have finished this thesis. I wish to extend my deep and warm gratitude to my father

Sadeq and my mother Zeinat. They raised me, taught me, and always supported and loved me. To them I dedicate this thesis.

In conclusion, I acknowledge that this research would not have been possible without the financial support of THALES, Australia through the ARC Linkage grant (LP0561200)/APAI Scholarship. To them, I express my sincere gratitude.

To my father and mother.

Contents

1	Introduction	1
1.1	Problem Statement	3
1.2	Thesis structure and contributions	6
1.3	Publications arising from this thesis	11
2	Literature Review	13
2.1	Wireless Sensor Networks	13
2.2	Sensor Faults, Drift, Bias and the Calibration Problem	16
2.3	Related Work	20
3	Drift Aware Wireless Sensor Networks	31
3.1	A Simple drift detection and correction algorithm	32
3.2	Evaluation	38
3.3	Conclusion	44
4	Correcting Measurement Errors under Smooth Drift Scenario	47
4.1	Smooth drifts estimation and measurements correction algorithm	48
4.2	Complexity analysis	55
4.3	Evaluation	55
4.4	Conclusion	60
5	Correcting Measurement Errors under Unsmooth Drift Scenario	63
5.1	Derivation of the IMM Algorithm	64
5.2	Unsmooth drifts estimation and measurements correction algorithm	71
5.3	Complexity analysis	77
5.4	Evaluation	78
5.5	Conclusion	85
6	Spatio-Temporal Modelling of Measurements in Wireless Sensor Networks	87
6.1	Modelling and predicting measurements using SVR	88
6.2	Iterative drift estimation and correction using SVR-KF framework	92
6.3	Complexity analysis	96
6.4	Evaluation	97
6.5	Conclusion	103

7	Addressing Estimation Errors Caused by Nonlinearity of SVR	105
7.1	Modelling and predicting measurements using SVR	106
7.2	Iterative measurement estimation and correction using an SVR-UKF framework	107
7.3	Complexity Analysis	113
7.4	Evaluation	114
7.5	Conclusion	125
8	Coping with Unsmooth Measurements and Under Sampled Data	127
8.1	Iterative measurement estimation and correction using SVR with UKF based IMM algorithm	128
8.2	Evaluation	131
8.3	Conclusion	136
9	Conclusions and Future Work	137
9.1	Conclusions	137
9.2	Future Research Directions	141

List of Figures

1.1	Wireless sensor area with encircled sub-network	3
1.2	Examples of smooth drifts	5
1.3	Examples of drifts with jumps and sudden changes	6
3.1	A comparison between NDASN and DASN under two time scenarios: a) NDASN, $DT = 50, 60, 70, 80, 90$ b) DASN, $DT = 50, 60, 70, 80, 90$, C) NDASN, $DT = 50$ d) DASN, $DT = 50$	39
3.2	Probability of network breakdown VS. Cluster size (n) for different drift scenarios and fixed communication channel reliabilities =1	40
3.3	Probability of network breakdown VS. Communication channel reliability for different drift scenarios and fixed cluster size $n = 10$	40
3.4	Breakdown time for different cluster sizes and different drift scenarios for $N = 100$ and $reliability = 1$: a) no drift, b) $a = 6, n = 1$,c) $a = 6, n = 5$,d) $a = 6, n = 10$	42
3.5	Breakdown time for different communication channel reliabilities and different drift scenarios for $N = 100, n = 10$: a) no drift, b) $a = 6, reliability = 0$, c) $a = 6, reliability = 0.3$, d) $a = 6, reliability = 1$	42
3.6	Probability of network breakdown VS. Cluster size for $N = (100, 50, 30)$, $reliability = 1$ and $a = 4.5$	43
3.7	Probability of network breakdown VS. Communication channel reliability for $N = (100, 50, 30)$, $n = 10$ and $a = 4.5$	43
4.1	A block diagram for the smooth drift estimation and measurement correction algorithm	55
4.2	Actual and estimated drifts in nodes 1 and 2 for when 2 sensors are drifting	57
4.3	The reading of node 1, the corrected reading and the actual temperature when 2 sensors are drifting	57
4.4	Actual and estimated drifts in nodes 1 and 2 when 7 sensors are drifting	58
4.5	The reading of node 1, the corrected reading and the actual temperature when 7 sensors are drifting	58
4.6	Actual and estimated biases/drifts in nodes 1 and 2	60
5.1	A block diagram for the unsmooth drift estimation and measurement correction algorithm	75
5.2	The IMM step, $\bar{\mu}_{i,k}^{\theta} = \mu_{k-1 k}^{\alpha \theta}$	76

5.3	The reading of node 1, the corrected reading and the actual temperature for KF.	79
5.4	Actual and estimated drifts in nodes 1 and 2 for KF.	79
5.5	The reading of node 1, the corrected reading and the actual temperature for IMM.	80
5.6	Actual and estimated drifts in nodes 1 and 2 for IMM.	80
5.7	Actual and estimated biases/drifts in nodes 1 and 2 for KF.	82
5.8	Actual and estimated biases/drifts in nodes 1 and 2 for IMM.	82
5.9	RMS error for both algorithms under smooth drift scenario.	84
5.10	RMS error for both algorithms under unsmooth drift scenario.	84
5.11	RMS error under unsmooth drift scenario for different number of models.	85
6.1	Support vector regression framework [91].	91
6.2	The SVR-KF drift estimation and measurement correction framework at node i	96
6.3	Sensor nodes in the IBRL deployment. Nodes are shown in black with their corresponding node-IDs. Node 0 is the gateway node [97].	98
6.4	Results for node ID 2 when only this node experiences a drift. The curves shown are (i) R-WD (ii) R-WOD (iii) DCM-WD (iv) DCM-WOD	101
6.5	Mean absolute error of readings for each scenario.	102
6.6	Mean absolute error of the corrected measurements for each scenario.	103
7.1	The SVR-UKF Measurement correction framework at node i	112
7.2	Results for node ID 2 when only this node experiences a drift. The curves shown are (i) R-WD (ii) R-WOD (iii) DCM-WD (iv) DCM-WOD.	118
7.3	Mean Absolute Error for the network without correction.	119
7.4	Mean Absolute Error for the network with correction for 2001 samples in 10 days	120
7.5	Mean Absolute Error for the network with correction for 4001 samples in 10 days	121
7.6	Estimated Drift in sensors with and without drift when the sampling rate is 2001 samples in 10 days.	123
7.7	Estimated Drift in sensors with and without drift when the sampling rate is 4001 samples in 10 days.	124
8.1	Measurement correction framework at node i for fast changing readings, $\overline{\mu}_{i,k}^{\theta} = \mu_{i,k-1 k}^{\alpha \theta}$	130
8.2	Mean Absolute Error for the network with correction for 2001 samples in 10 days using 11 levels IMM.	132
8.3	Mean Absolute Error for the network with correction for 2001 samples in 10 days using 7 levels IMM.	134
8.4	Mean Absolute Error for the network with correction for 2001 samples in 10 days using 5 levels IMM.	134
8.5	Mean Absolute Error for the network with correction for 2001 samples in 10 days using 3 levels IMM.	135

List of Tables

5.1	Processing times required by KF based and IMM based drift estimation and correction algorithms.	85
7.1	Sensor nodes IDs, their assigned neighbours and the SVR parameters (C and γ_C) for Case 2 with 2 sampling rates.	116
7.2	Correlation Coefficients of Node ID 32 with it's neighbours at the training phase ρ_t and running phase ρ_r	125
8.1	Processing times required by SVR-UKF based and IMM-SVR-UKF based error correction algorithms.	133
8.2	Energy consumed for each sensor action, based on measurements of the Mica2 sensor node qouted from [108].	135

List of Abbreviations

DCM-WD	Drift Corrected Measurement With Drift
DCM-WOD	Drift Corrected Measurement Without Drift
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
FFT	Fast Fourier Transform
KF	Kalman Filter
IBRL	Intel Berkeley Research Laboratory
IMM	Interacting Multiple Model
R-WD	Reading With Drift
R-WOD	Reading Without Drift
SVM	Support Vector Machine
SVR	Support Vector Regression
UKF	Unscented Kalman Filter
UT	Unscented Transform
WSN	Wireless Sensor Network

Abstract

WIRELESS Sensor Networks (WSNs) are deployed for the purpose of monitoring an area of interest. Even when the sensors are properly calibrated at the time of deployment, they develop drift in their readings leading to erroneous network inferences. Traditionally, such errors are corrected by site visits where the sensors are calibrated against an accurately calibrated sensor. For large scale sensor networks, the process is manually intensive and economically infeasible. This imposes finding automatic procedures for continuous calibration. Noting that a physical phenomenon in a certain area follows some spatio-temporal correlation, we assume that the sensors readings in that area are correlated. We also assume that measurement errors due to faulty equipment are likely to be uncorrelated. Based on these assumptions, we follow a Bayesian framework to solve the drift and bias problem in WSNs.

In the case of densely deployed WSN, neighbouring sensors are assumed to be close to each other that they observe the same phenomenon. Hence, the average of their corrected readings is taken as a basis for each sensor to self-assess its measurement, estimate its drift and to correct the measurement using a Kalman Filter (KF) in the case of smooth drift, and the Interacting Multiple Model algorithm (IMM) in the case of unsmooth drift. The solutions are computationally simple, decentralised and also scalable. Any new node joining the neighbourhood needs only to obtain the corrected readings of its neighbours to find the average and apply the KF iterative procedure.

On the other hand, when the sensors are not densely deployed, Support Vector Regression (SVR) is used to model the interrelationships of sensor measurements

in a neighbourhood. This enables the incorporation of the spatio-temporal correlation of neighbouring sensors, to predict future measurements. The SVR predicted value is used by a KF to estimate the actual drift and correct the measurement. Unfortunately, the KF introduces some system errors when used with nonlinear systems. The use of Unscented Kalman filter (UKF) instead, considerably reduces the system error and results in a better drift correction. The use of IMM with the SVR-UKF framework allows for reducing the sampling rate which eventually reduces the communication overhead among the sensors and saves the communication energy.

In this thesis, we present several solutions for the random and systematic (drift and bias) errors in sensors measurements, for different sensor deployment scenarios. We also consider two drift scenarios, namely smooth and unsmooth drifts. We evaluate the presented algorithms on simulated and real data obtained from the Intel Berkeley Research Laboratory sensor deployment. The results show that our algorithms successfully detect and correct systematic errors (drift and bias) developed in sensors and filters out the noise. Thereby, prolonging the effective lifetime of the network.